Project Report

Constructor University

MSc. Data Science for Society and Business

Polypropylene:
price drivers and forecast

Authors:

Zerong Lu (30008217)

Trifanov Matvei (30008383)

Perera Pasindu (30008137)

Supervisor: Prof. Dr. Christoph Lattemann

Module Name: Digital Transformation of Organizations

Date of submission: <29 12 2025>

0 Table of Contents

1 Introduction

Polypropylene, or PP, stands as one of the most common plastics in daily life. People use it for packaging, car parts, textiles, and many household items. This material comes from oil through a clear process: crude oil turns into polymer-grade propylene, or PGP, and then into PP. Because of this link, PP prices tend to follow changes in energy prices. But the match is not perfect. PP often shows smaller ups and downs compared to crude oil. Other factors, like supply issues or demand shifts, play a role too. Big events or news can sway prices in ways that basic energy trends miss. This paper explores macro news to find those extra clues.

The topic holds real value today. Oil price swings have hit hard in recent years. Think of the 2020 price war between big producers or the energy crunch in 2021-2022 from global events. These shocks raise costs for companies that rely on PP. Better forecasts can help firms manage risks and plan ahead. Accurate predictions matter for supply chains and budgets. This paper taps into OPEC Monthly Oil Market Reports from 2019 to 2025. These reports offer steady, reliable views on oil trends. Tools like GPT and FinBERT pull sentiment from the text to build a hybrid index. The index looks at parts on demand, supply, and price outlooks.

Sentiment analysis adds a layer to traditional price models. It captures the tone of market narratives, which might signal future moves. The approach combines data from prices with text insights. This helps test if news tones explain PP changes beyond energy basics. The research questions guide the work:

How much do PP prices move with crude oil and other energy stuff?

- What part of PP changes can't crude explain alone?
- Does OPEC sentiment link to those leftover PP parts?
- Does sentiment make short-term PP forecasts better?

This paper uses time series tools, correlations, and models to check these points. It draws on monthly data from Europe and global sources. Charts show trends, volatility, and links. The goal is to see if adding text data sharpens forecasts. In the end, the findings shed light on PP drivers in a changing energy world.

2 Data Description and Sources

This paper relies on price data for PP, PGP, and energy items like crude oil. It also includes text from OPEC reports to build sentiment scores. All data goes through cleaning steps. Weekly or daily figures get turned into monthly averages. The sets join up for a shared timeline of 123 months, from March 2015 to October 2025. Code in the repo handles this. It scans files and spots gaps, like missing naphtha data. The main variables—PP, PGP, and crude—have no missing points in the joined sample.

PP prices focus on Europe, in euros per ton. Sources include files like polypropylene_primary_avg_prices.csv. This gives monthly averages from August 2018 to April 2025, with 81 rows. Other files, such as polypropylene_weekly.csv and polypropylene_weekly_clean.csv, provide weekly prices from January 2015 to March 2025, each with 520 rows. These capture spot and average market values. PGP data follows a similar pattern. It uses the same files for monthly and weekly views. PGP acts as the key input for PP making.

Crude oil data comes from benchmarks. Files like crude_oil_daily.csv offer daily closes from January 2015 to October 2025, cleaned to 2,712 rows. Weekly versions include crude_oil_weekly.csv with 564 rows and crude_oil_weekly_clean.csv with 1,128 rows. The paper adds related energy items: Brent, WTI, and natural gas. These help check wider links. All prices get rebased or normalized for fair comparisons. For example, levels start at 100 in 2015 to show trends clearly.

Text data comes from OPEC Monthly Oil Market Reports. These full texts span 2019 to 2025. FinBERT scores them for positive or negative tone. GPT adds context by comparing tones across reports. The hybrid index starts at zero in January 2019. It tracks changes over time, like shifts in demand or supply views. The repo notes extra text from BVSE plastics reports, but OPEC stays the main focus.

Python code from the GitHub repo manages everything. The link is https://github.com/intox1ca7ed/LLM-polypropylene. Scripts use pandas to clean and join data. They make log returns for stable analysis and residuals to spot unique PP moves. Plots and cleaned files sit in the repo for easy checks. This setup ensures the data is solid and ready for models.

3 Exploratory Data Analysis of Prices

This section examines the price behavior of polypropylene (PP) in relation to key factors affecting the raw materials and energy markets, with an emphasis on two issues. First, do PP prices move in the long term along with the prices of raw materials and oil? Secondly, does the short-term relationship provide sufficient grounds for forecasting, or does it leave significant unexplained fluctuations that require additional sources of information, presented later in the report. The preliminary analysis uses both price levels (to understand general trends and major shifts in the regime) and monthly logarithmic returns (to assess the short-term relationship and forecast significance).

3.1 Price levels: co-movement with divergence in magnitude

We begin with comparing the price levels for polypropylene (PP), polymer propylene (PGP) and crude oil after bringing each batch to a common baseline (base = 100). The normalized view highlights that PP and its underlying factors are subject to similar broad cycles, including major market events such as the 2020 recession and the strong upswing in 2021-2022. At the same time, PP tends to adjust more slowly and with a smaller amplitude than crude oil, reflecting the fact that PP is an industrial material, the price of which depends on contractual structures, inventory dynamics and supply chain friction. In other words, the graph of levels corresponds to the economic intuition that crude oil and propylene are a strong anchor of costs, but do not fully determine the trajectory of PP from month to month.

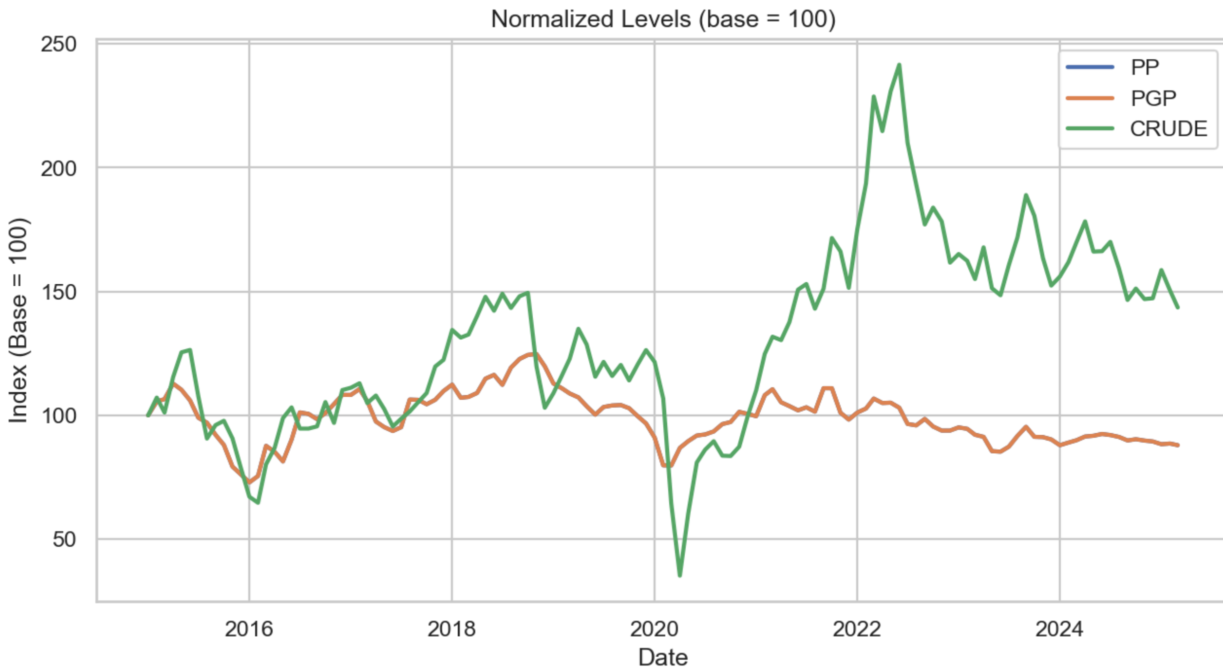

Fig 3.1 : Normalized price levels (base = 100): PP vs PGP vs Crude

The key methodological point is that the joint movement of levels can overestimate short-term predictability. General trends (inflation, global demand cycles, and major general shocks) mechanically increase level correlations. For forecasting and testing incremental characteristics, the relevant question is whether changes in the oil markets help explain the changes in PP over the monthly horizon. For this reason, the rest of the EDA focuses on profitability and the diagnosis of time relationships.

3.2 Monthly returns: a modest contemporaneous link

To quantify the short-term joint dynamics, we calculate the monthly logarithmic returns of PP and crude oil and visualize this relationship using a scatter plot with an OLS approximation line. The yield dispersion shows significant variance around the approximation line: in many months, PP movements do not coincide with crude oil movements in the same month, consistent with the idea that PP has an idiosyncratic component that goes beyond crude oil.

In the sample used for this EDA, the current correlation between PP and crude oil yields is 0.21, which is a positive but not a strong indicator. This result is important because it defines crude oil as an informative but insufficient independent predictor of PP movement for one month ahead.

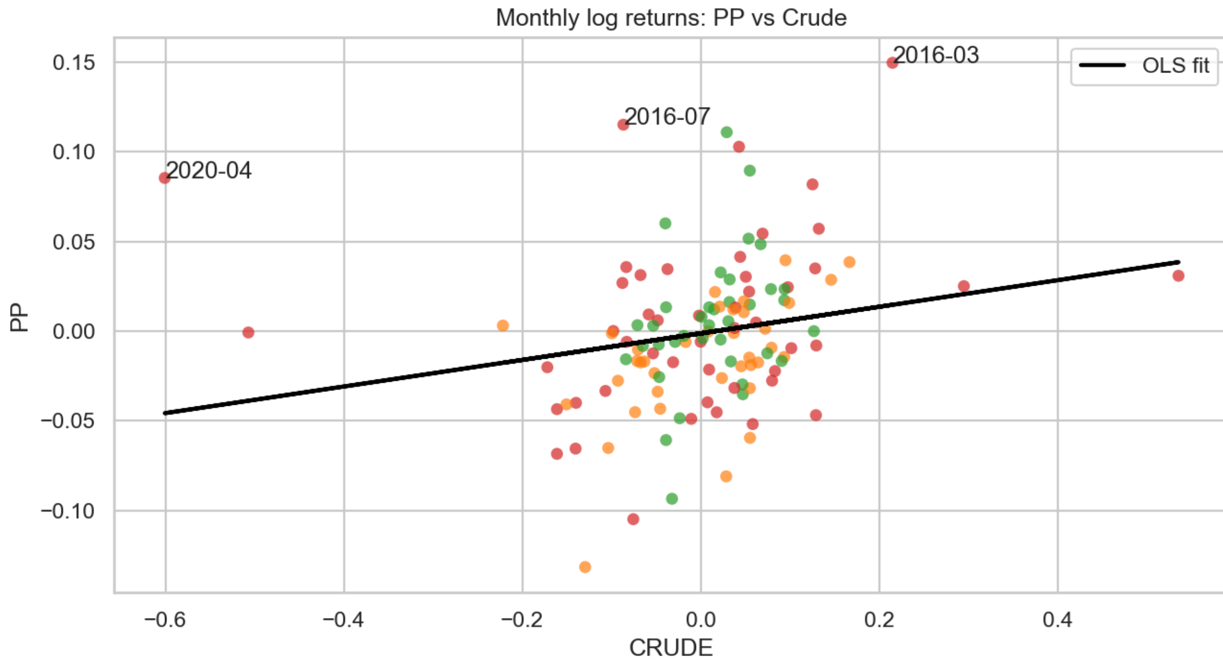

Figure 3.2: Dispersion of monthly logarithmic returns: PP compared to crude oil

In the sample used for this EDA, the current correlation between PP and crude oil yields is 0.21, which is a positive but not a strong indicator. This result is important because it defines crude oil as an informative but insufficient independent predictor of PP movement for one month ahead. There is a short-term relationship, but it is weak or moderate; large unexplained scattering motivates residue-based modeling.

3.3 Time variation and timing: rolling beta and lead–lag structure

The relationship between commodities is not constant over time. Changes in market structure, supply disruptions, changes in refinery margins, and political events may affect the degree to which refined product prices depend on oil prices. To account for this, we calculate rolling 12-month estimates of PP sensitivity to the oil price (beta coefficient) and the corresponding moving R-squared coefficient from the same moving regression. Rolling diagnostics show that this relationship depends on the regime: there are periods when crude oil explains PP fluctuations to a greater extent, and periods when compliance worsens. This directly supports a modeling approach that takes instability into account, and helps explain why a fixed relationship based on a complete sample can produce inaccurate results in certain years.

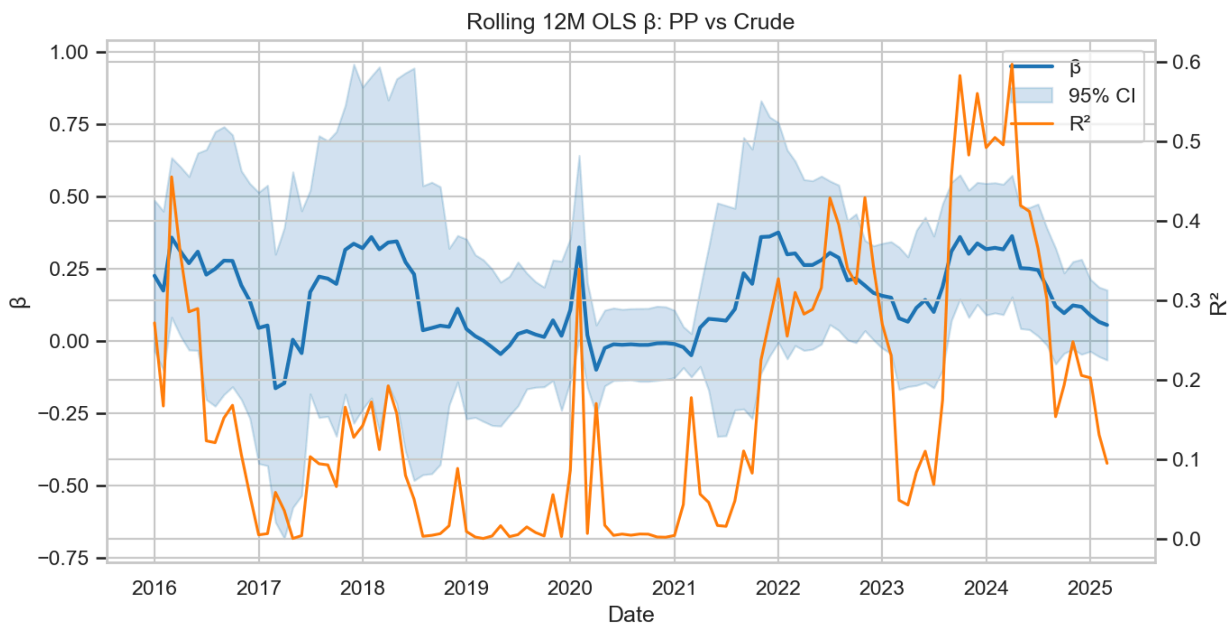

Figure 3.3.1: Rolling 12M beta and rolling R-squared for PP on crude returns

We also check whether PP reacts with delay to crude oil shocks using advance/delay correlation scanning. The advance/lag analysis shows that the greatest overlap is observed when crude oil is ahead of PP by about one month: the best advance/lag configuration is -1 lag with a correlation of 0.36 (which is interpreted as crude oil being ahead of PP by one month). Although this result is still modest in magnitude, it is useful from a practical point of view, as it justifies focusing on the lagging variables of crude oil in later baseline regressions and forecasting specifications.

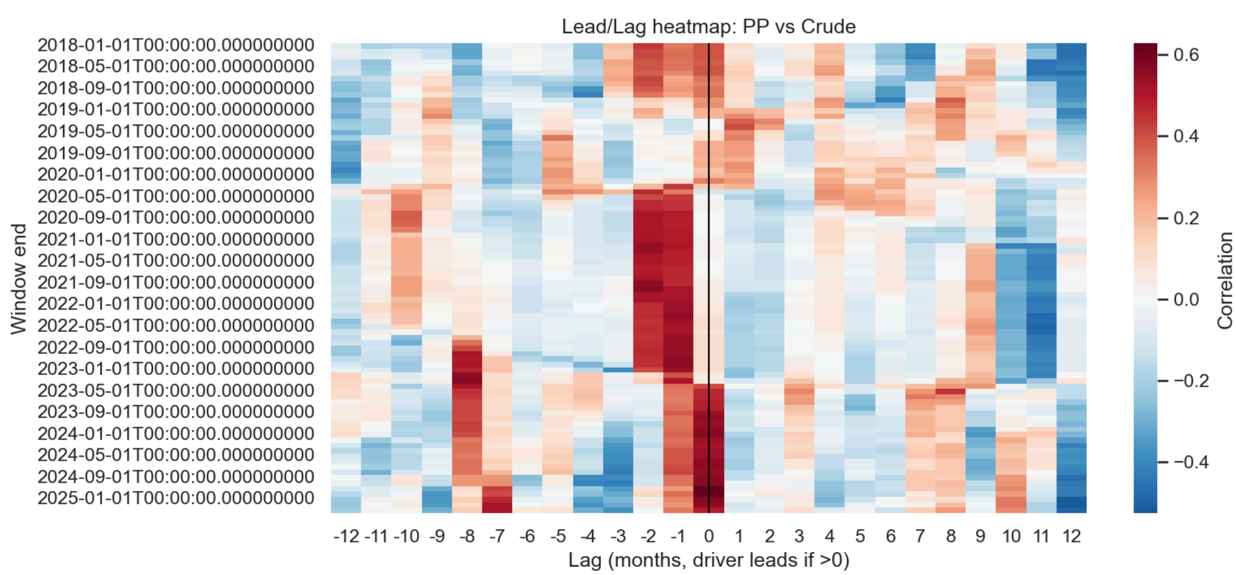

Figure 3.3.2: Lead/lag correlation of PP vs crude monthly returns.

Again, the maximum correlation occurs when crude leads by ~1 month, motivating the use of lagged crude features.

3.4 The broader energy market context: correlations between benchmarks

Finally, we consider PP in the broader context of the energy complex, comparing its relationship with the main benchmarks for crude oil prices and related energy variables.

	mean	std	min	max
PP	-0.001	0.04	-0.131	0.14
PGP	-0.001	0.04	-0.131	0.14
CRUDE	0.002	0.00	0.12	0.53

Table 4.4.1 Summary statistics of monthly log returns

Correlation analysis shows that reference oil prices such as WTI and Brent show a relatively strong positive relationship with PP in a broader correlation diagnosis (for example, a correlation of about 0.70 for WTI and 0.67 for Brent in the presented lag correlation results), consistent with the cost relationship between oil markets and polymer pricing in in the field of processing and production.

PP_EU	1.0
Brent_Lag1	0.674
WTI_Lag1	0.696
NatGas_Lag1	0.621

Table 3.4.2 The correlation matrix of monthly logarithmic returns

However, the analysis also confirms that correlation models depend on whether levels or returns are measured, as well as on the lag structure. This served as motivation for the modeling presented later in the report: using market variables (oil and related factors) as basic characteristics, while recognizing that a significant part of PP fluctuations remains unexplained over the monthly horizon.

Overall, the preliminary results of the study confirm the three conclusions that underpin the rest of the report. First, in the long term, PP moves synchronously with energy markets, which is consistent with cost shifting and general macroeconomic shocks. Secondly, in the monthly horizon, the current relationship between PP and crude oil yields is moderate (0.21), and even the best advance/lag match remains limited (0.36 with one month ahead of crude oil), indicating significant residual variation. Thirdly, this relationship is unstable over time, which leads to the formation of a model based on residual values and, subsequently, to a thorough check of whether the sentiment derived from the text adds additional information beyond the standard market variables.

4 Baseline Price Modeling and Residual Construction

This section sets out two foundations for the rest of the project. First, we define simple basic models to establish a realistic benchmark of performance and interpretability for polypropylene (PP) price behavior. Secondly, we construct an idiosyncratic series of residual PP values by removing the component of PP fluctuations that can be explained by the yield of crude oil. This balance is a target signal that is expected to be explained later using sentiment characteristics, as it isolates the part of PP fluctuations that is not accounted for by the main cost-influencing factor in the extractive industry

4.1 Basic forecasts of the price level

We'll start with simple basic forecasts based on price levels to see how far we can go with direct extrapolation. The evaluation uses a time series: the last part of the sample is reserved as a test period, which is approximately 20% of the available observations and is limited to 24 months. We then compare the observed PP prices in the test period with the baseline forecasts obtained using (i) a naive forecast that projects the latest observed value forward, (ii) a seasonal naive benchmark, where applicable, and (iii) a Holt-Winters exponential smoothing model to account for trend and seasonality. The goal is not to "win" using complex modeling, but to quantify how much of the PP dynamics can only be explained by consistency and a smooth seasonal structure. The visual overlap of predicted and actual levels allows you to intuitively check whether these simple methods track turning points and major regime shifts.

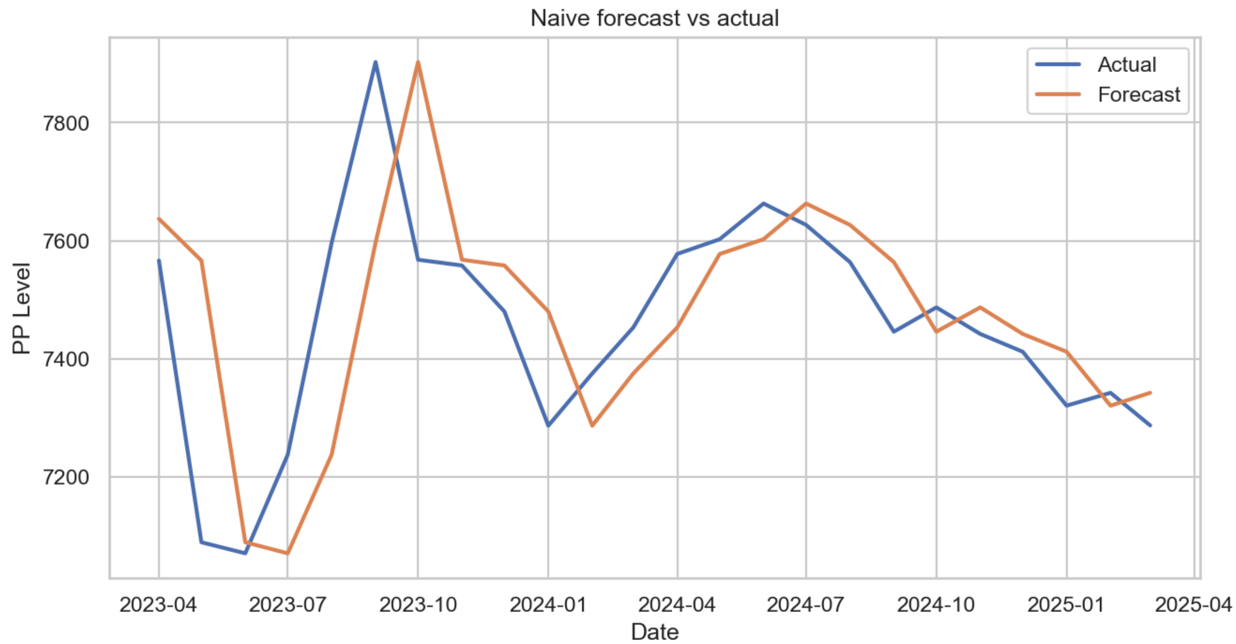

Figure 4.1.1 PP level baselines: actual vs predicted

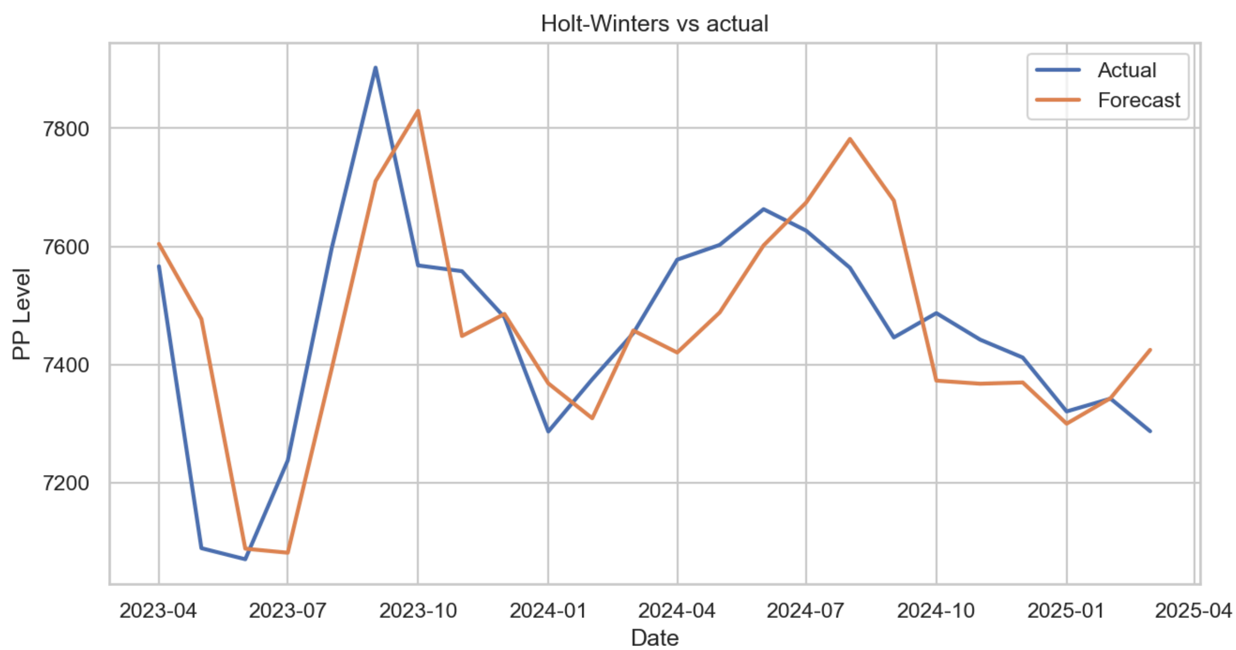

Figure 4.1.2 Holt–Winters vs actual

4.2 The basic level of profitability determined by oil prices

Since PP is related to energy markets, we also build a base level of profitability determined by oil prices. Monthly logarithmic PP yields are regressed by lagging oil yields (using a three-month lag). This specification reflects the hypothesis that changes in oil prices can be transmitted to PP with a delay, consistent with frictional effects, contract reductions, inventory adjustments, and delivery times. The regression is intentionally simplified in order to preserve the interpretability of the coefficients and simplify subsequent comparisons with models supplemented by sentiment.

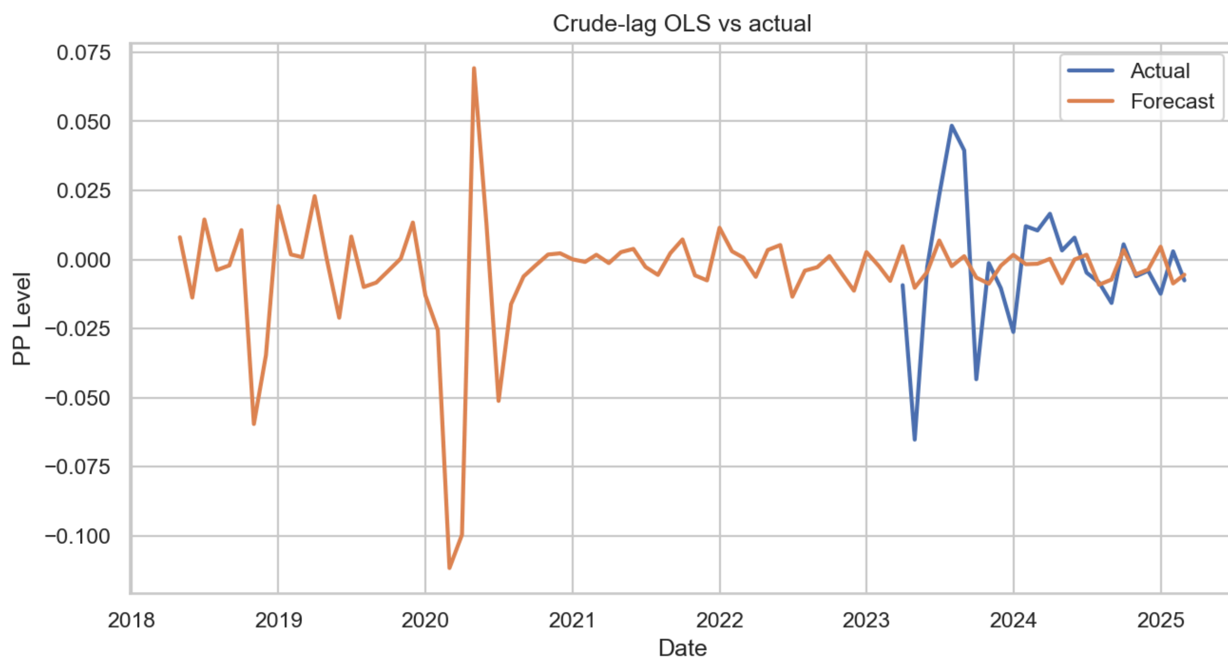

Figure 4.2 Crude-driven regression

To avoid overly optimistic sample-based conclusions, we evaluate the regression using a rolling-onset past test. At each stage, the model is reconfigured based on the available history and used to predict the next observation, simulating a realistic forecasting situation where parameters are estimated based only on past data. In the observed run, the oil price lag model reaches a low explanatory power with an R2 of about 0.05, indicating that oil yields explain only a small part of the monthly fluctuation in PP yields. At the same time, at least one lagging oil price term is statistically significant ($p<0.05$), which supports the idea that the oil price contains some predictive information, but it is not sufficient to fully characterize the PP movement over a one-month horizon.

4.3 Construction of an idiosyncratic remainder PP

The main result of this section is the idiosyncratic remainder of PP. After fitting the delayed crude oil regression to the aligned yield data, we calculate the fitted values (the component explained by crude oil) and define the remainder as the difference between the actual PP yield and the regression forecast. By design, this series of residues is centered around zero and represents the "innovations" of PP that remain after accounting for the cost channel associated with crude oil. Conceptually, it reflects factors related to refining and the market such as changes in polymer production capacity, regional logistical constraints, contract terms, inventory cycles, and demand shocks that are not directly reflected by crude oil yields.

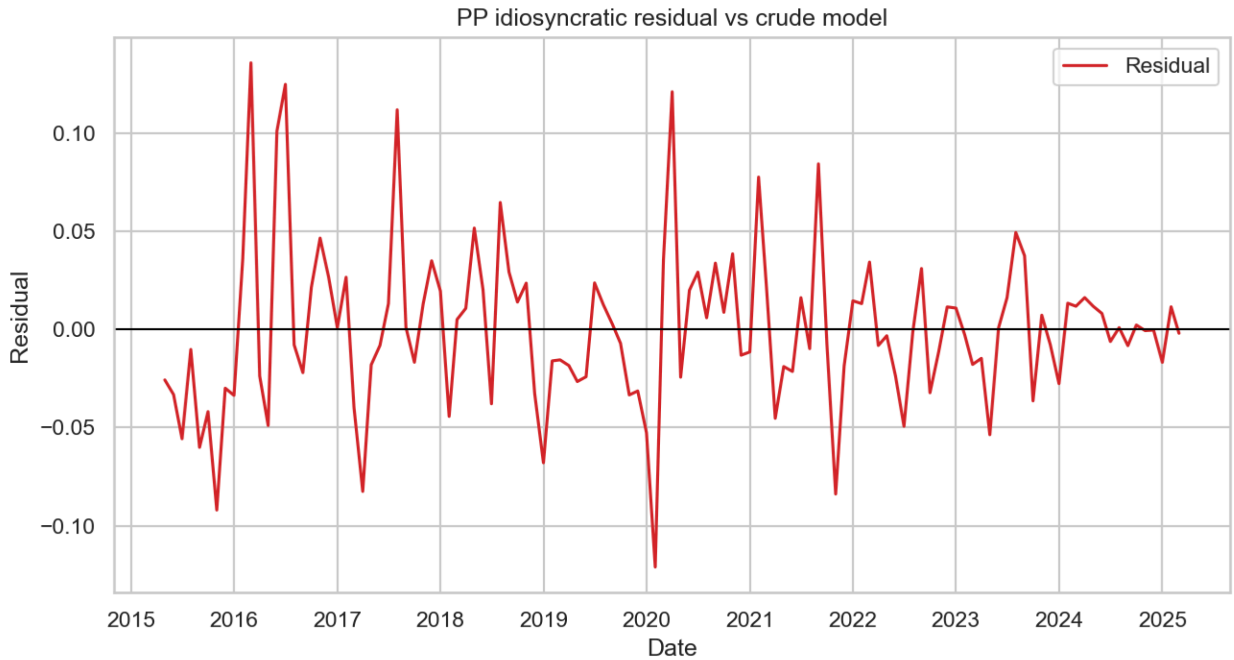

Figure 4.3 Idiosyncratic residual (PP residual = PP return − crude-fitted return)

A graph of the time series of residual values is useful for interpretation: periods of sustained positive or negative residual values correspond to episodes when PP consistently exceeds or is inferior to what crude oil assumes. This makes the residual values a suitable target for checking whether the OPEC report provides additional information: if the report contains real forward-looking signals about market tensions, demand prospects, or price expectations, it should correlate with or help predict these residual deviations.

4.4 Regime context and diagnostics

Two additional diagnostics help interpret the residual and justify later modeling choices. First, a decomposition of PP levels (e.g., STL) can separate long-run trend, seasonal structure, and irregular components; this is only worth keeping if the report explicitly explains what the decomposition reveals about trend/seasonality versus short-term shocks. Second, a rolling volatility comparison between PP and crude provides context on changing risk regimes: in high-volatility periods, even a stable structural relationship can appear weaker, and forecast errors can increase regardless of model choice.

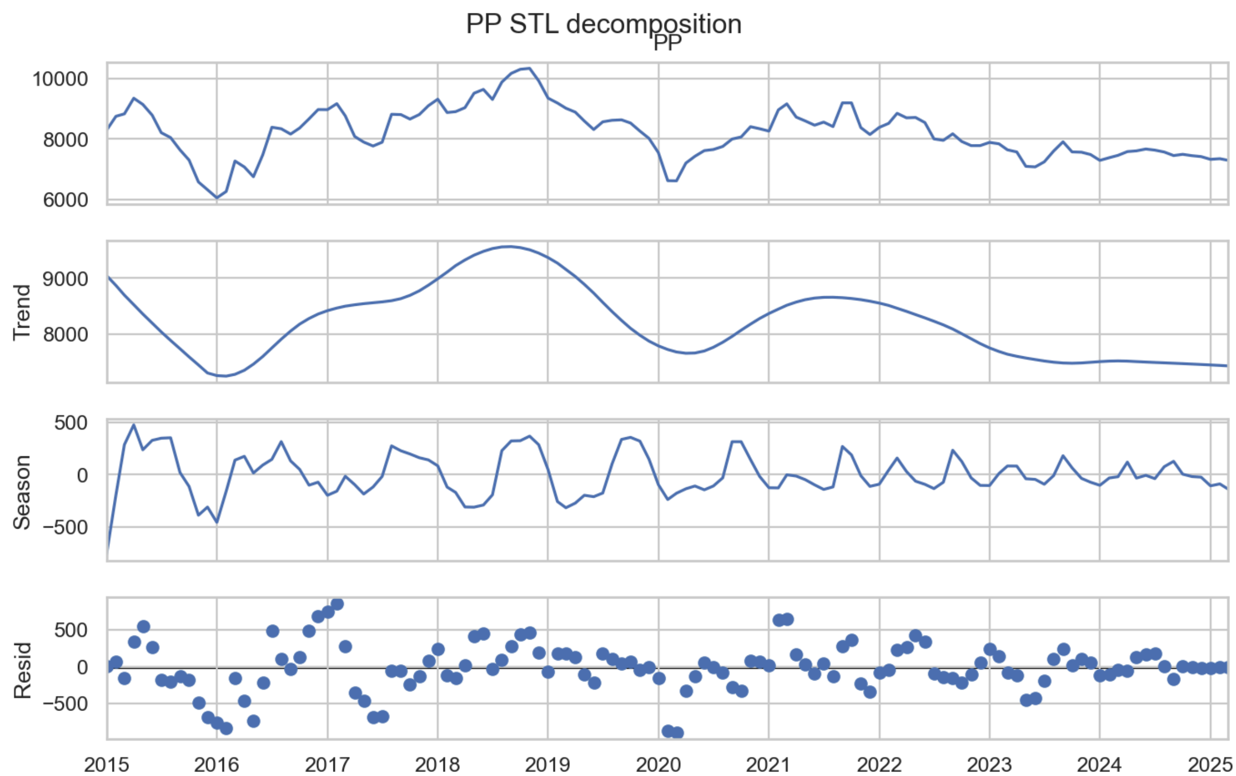

Figure 4.4.1 STL decomposition of PP price level

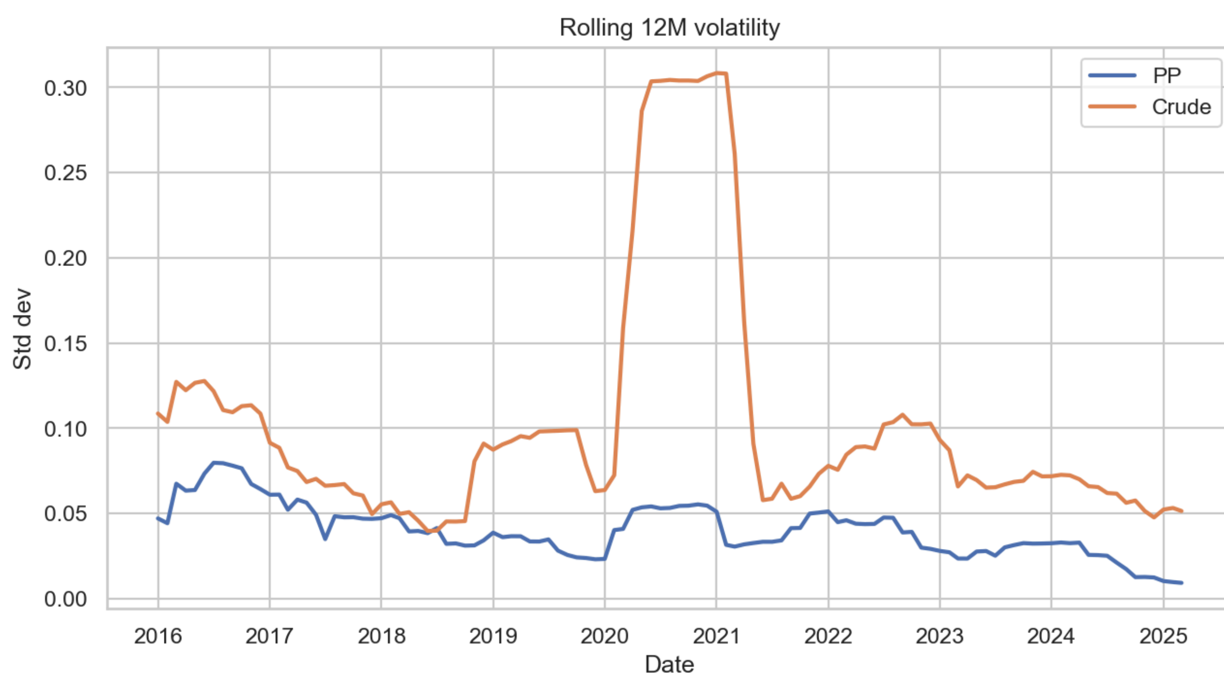

Figure 4.4.2 Rolling volatility: PP vs crude returns

Suggested tables and figures for this section are therefore: a baseline forecast overlay plot (levels), the crude-lag OLS backtest plot (returns), and the PP residual time series around zero. If you include tables, keep them minimal and consistent in scale: one table for level-baseline errors in level units, and if needed, a separate short table for return-model diagnostics (R^2 and significant lags), rather than mixing level and return metrics in a single scorecard.

5 Text Processing and Sentiment Analysis

5.1 OPEC Report Preprocessing and Sectioning

The lengthy, unstructured OPEC Monthly Oil Market Reports (MOMR) include a variety of information, from comprehensive supply and demand statistics to macroeconomic commentary. Treating the entire report as a single text unit is neither efficient nor theoretically sound for sentiment-based analysis because different sections frequently convey different and sometimes opposing signals.

In order to address this, every OPEC report was methodically preprocessed and divided into sections that were economically significant, with an emphasis on those that were most pertinent to the formation of commodity prices. The sections listed below were taken out and examined independently.

Section	Content	Relevance
Demand Outlook	Global oil demand, consumption and economic growth	Demand side impact on the price
Supply Outlook	OPEC and Non - OPEC production, outages and spare capacity	This sections reflects the over/under supply of oil that influence the oil prices
Overall price outlook and future demand supply projections	OPEC’s qualitative assessments of oil prices and demand and supply projections over the coming months	Direct forward guidance on the price expectations
Overall Report	Combination of all the sections	Captures the holistic market sentiment.

These ectioning allows the analysis to **disentangle demand-driven sentiment from supply-driven sentiment**, which is critical when studying commodities where price movements often result from asymmetric shocks.

5.2 FinBERT Sentiment Analysis: Motivation and Limitations

Reasons to Use FinBERT

Initially, each OPEC report section's sentiment was classified using FinBERT, a domain-specific transformer model trained on financial text. FinBERT was used for three reasons:

- It was created especially for financial terminology.
- It offers repeatable sentiment scores (neutral, negative, and positive).
- Large-scale automated processing without manual labeling is made possible by it.

Each report section's FinBERT sentiment score was calculated independently, and the results were combined monthly.

FinBERT's Drawbacks in Commodity Reports

When used on OPEC reports, FinBERT shows significant limitations despite its financial training:

- Context insensitivity: FinBERT frequently categorizes statements like "supply growth slowed" as neutral or negative because of their superficial language, even though they may be economically bullish.
- Narrative complexity: FinBERT finds it difficult to interpret OPEC reports holistically because they often balance opposing forces, such as tighter supply but weaker demand.
- FinBERT captures tone rather than economic implications, demonstrating a lack of economic reasoning.

In order to supplement FinBERT, a GPT-based sentiment approach was integrated due to these limitations.

5.3 GPT-Based Comparative Sentiment Approach

A GPT-based large language model (LLM) was used to conduct comparative, economically grounded sentiment analysis in order to get around FinBERT's drawbacks.

GPT, in contrast to FinBERT, is directed to function as a commodity market analyst, specifically reasoning about:

- Balance between supply and demand
- Availability of feedstock
- Implications for the macroeconomy
- Price pressure in the future

GPT was asked to do the following for every report section:

- Analyze the text's economic meaning.
- Determine whether the sentiment is neutral, bullish, or bearish.
- Give a confidence score that represents the signal's strength.

This method is more appropriate for complex macroeconomic narratives since it allows semantic understanding instead of keyword-based polarity detection.

5.4 Construction of the Hybrid Sentiment Index

A Hybrid Sentiment Index (HSI) was created to combine FinBERT and GPT because of their complementary strengths.

Logic for Hybrid Indexes

- FinBERT offers scalability and consistency.
- GPT offers contextual awareness and economic reasoning.

5.5 Baseline Normalization and Month-over-Month Changes

Normalization at Baseline

The Hybrid Sentiment Index was normalized using January 2019 as the baseline to enable intuitive interpretation over time. This standardization guarantees that:

- In comparison to the baseline, positive values show a growing sense of optimism.
- An increasingly pessimistic attitude is indicated by negative values.
- Month-to-Month Variations

Month-over-month changes were calculated in order to capture short-term sentiment shocks: Sentiment inflection points, which are especially important for event-study-style analysis, can be found in this way.

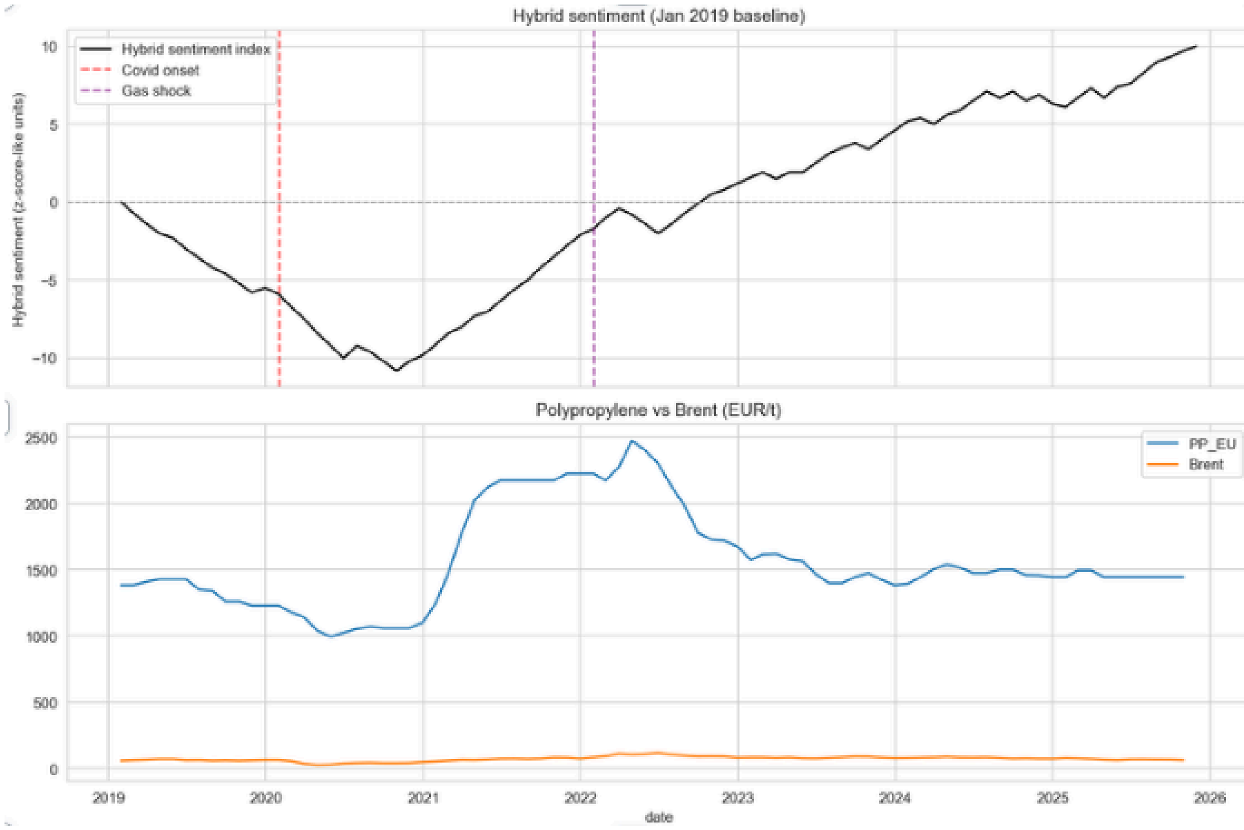

Fig 5.5 : Baseline Normalization of OPEC report

5.6 Empirical Results and Visual Analysis

Hybrid Sentiment Index over time:

There are distinct structural phases visible in the normalized Hybrid Sentiment Index:

- Sharp decline in sentiment during disruptions to the global economy.
- Phases of recovery coincided with tightening supply and increases in demand.
- Times when different sections' sentiments diverge.

Changes in Sentiment from Month to Month:

The month-over-month change bar graph illustrates:

- Unexpected emotional shocks brought on by macroeconomic or geopolitical developments.

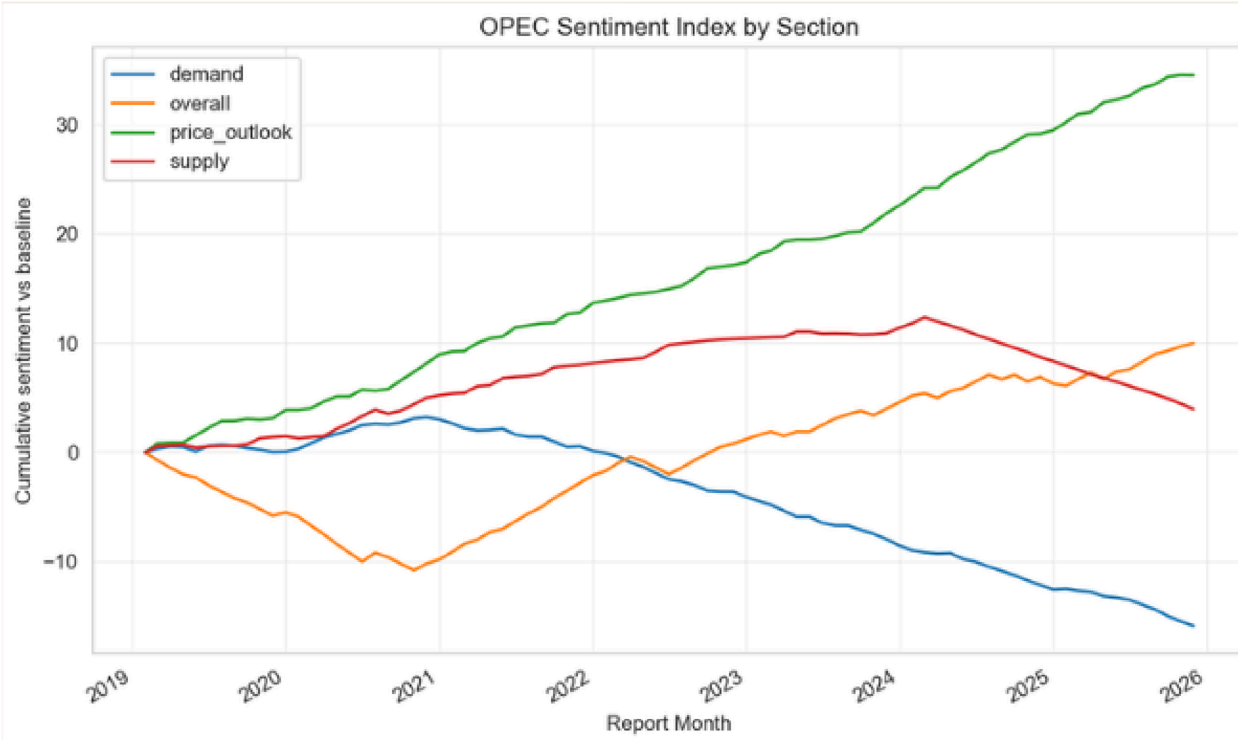

Fig 5.6.1: Hybrid Sentiment: Month-to-Month Shifts (Shock Detection)

- Leading indicators of price volatility as opposed to price levels.

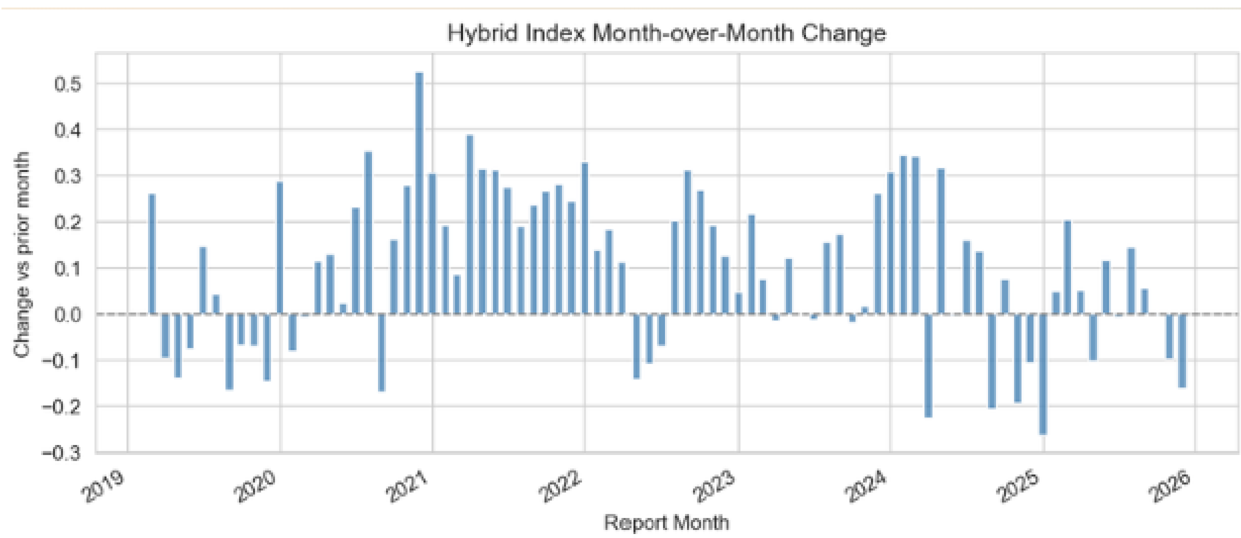

Fig 5.6.2: Hybrid Index MoM Change

6 Sentiment–Price Relationship Analysis

This section looks into whether the created Hybrid Sentiment Index has information that is economically significant for predicting or explaining changes in the price of polypropylene (PP). To isolate informational effects beyond basic price trends, the analysis specifically focuses on the relationship between sentiment and PP price residuals rather than raw prices.

6.1 Methodological Rationale

Long-term demand trends, feedstock costs, and crude oil prices are some of the structural fundamentals that significantly influence polypropylene prices. PP prices were first detrended by regressing them on important fundamentals in order to prevent spurious correlations. For evaluating the incremental explanatory power of sentiment, the resulting residuals, which show unexplained price deviations, are more appropriate.

6.2 Time-Series Relationship Between Sentiment and PP Residuals

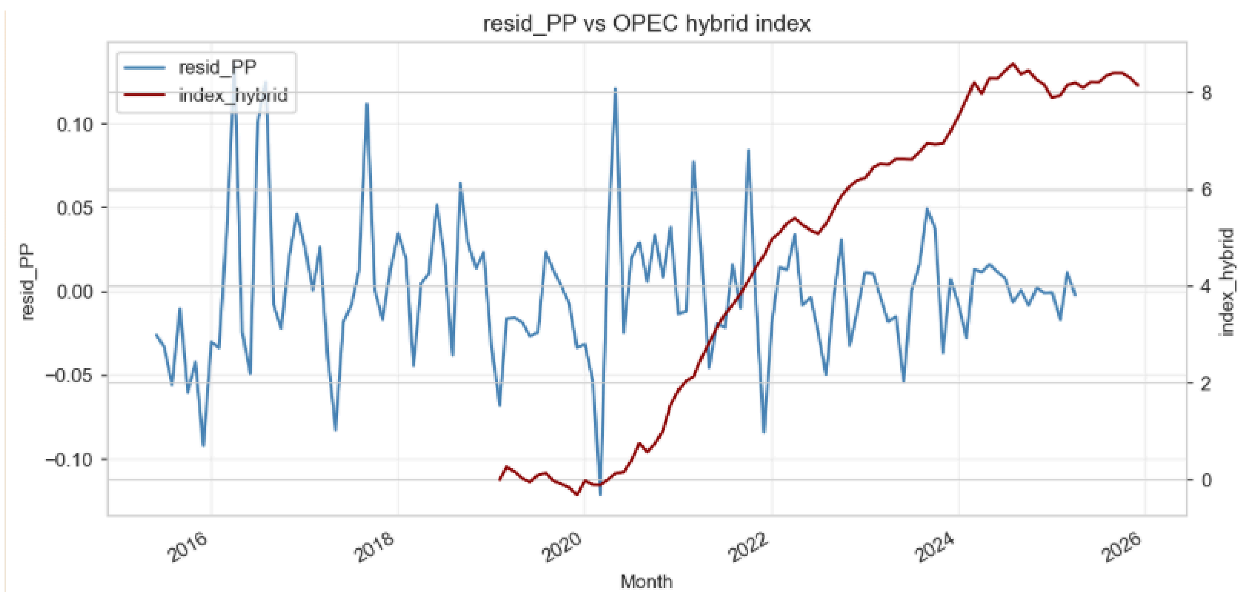

Fig 6.2 :resid_PP vs OPEC hybrid index – dual axis time series
A dual-axis time-series comparison of the OPEC Hybrid Sentiment Index and PP price residuals is shown in the first figure.

From 2019 onward, the hybrid sentiment index shows a significant upward trend, which is consistent with OPEC reports' increasingly positive or encouraging narratives during the post-pandemic recovery and ensuing market tightening. PP residuals, on the other hand, show transient positive and negative deviations and oscillate around zero without a consistent trend.The two series do not exhibit consistent long-term co-movement visually. Nonetheless, some brief episodes demonstrate that sentiment shifts come before transient changes in PP residuals. Residuals usually quickly return to zero after these episodes, indicating that sentiment effects are fleeting rather than long-lasting.

Although sentiment may occasionally affect short-term deviations, this time-series comparison already shows that sentiment does not drive PP prices structurally.

6.3 Scatter Diagnostics: Strength of the Sentiment Signal

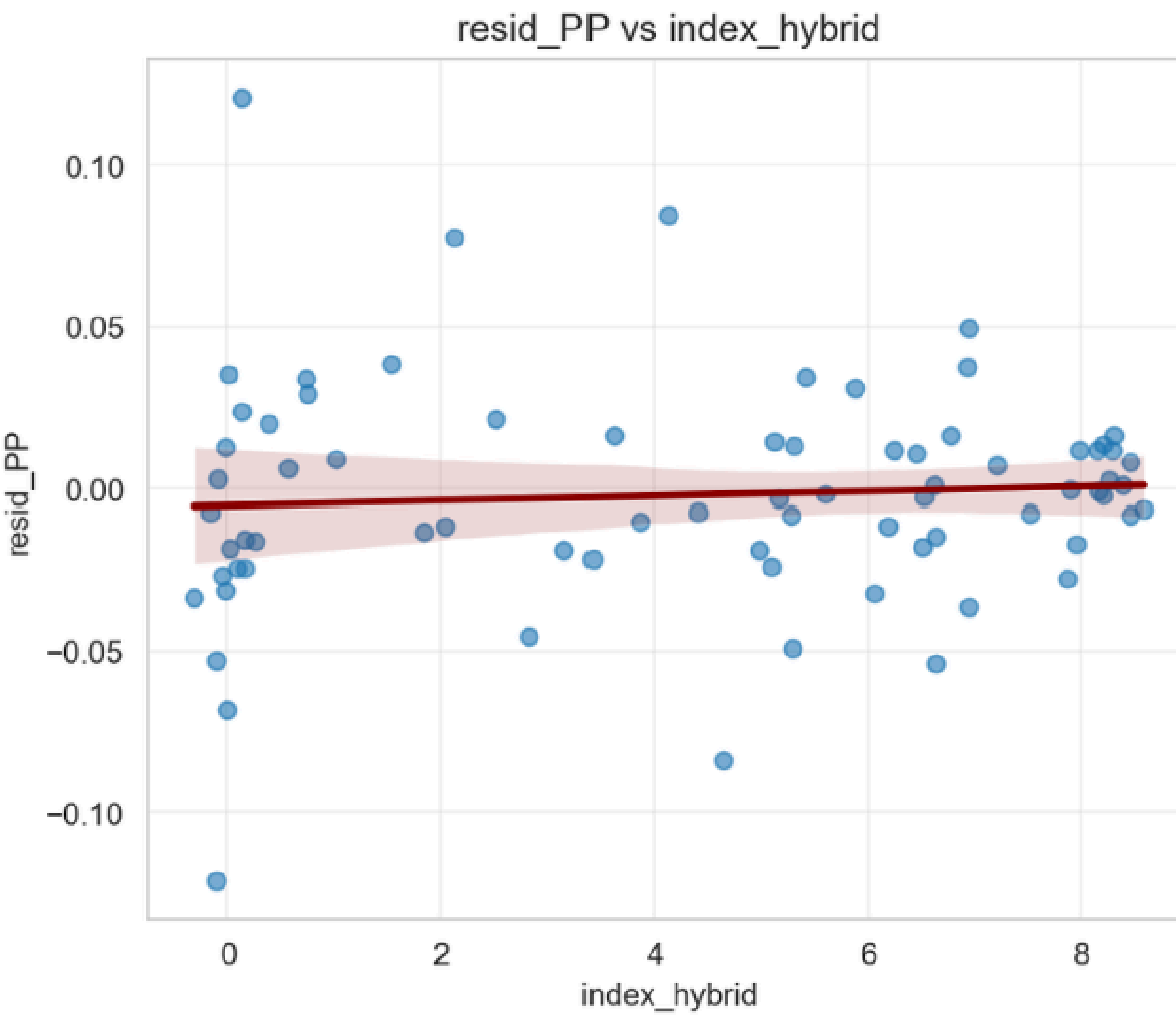

Fig 6.3: Scatter plot: resid_PP vs index_hybrid with regression line
In the scatter plot, the PP residuals are directly plotted against the Hybrid Sentiment Index values.

The linear regression reveals a slightly positive linear relationship, suggesting that higher values of sentiment are associated with higher PP residuals, but the scatter of points around the regression line is large, and the explanatory power of sentiment can be considered clearly limited.

This graph verifies that:

- The sentiment signal is present, but
- Its strength is small, and
- It is not sufficient to explain most of the variation in PP price residuals.

Practically, the meaning of sentiment might be understood as something secondary or auxiliary, rather than something that explains alone.

7 Forecasting Framework

This section describes the forecasting setup used to check whether the OPEC report adds additional predictive value to European Polypropylene (PP). An earlier analysis showed that PP is related to energy markets, but that oil itself leaves a significant unexplained component. To focus on this "unexplained" part, we forecast a residual PP series for the next month (resid_PP), which represents a part of the PP movement that is not covered by the baseline scenario based on the oil price. This formulation of the problem turns forecasting into a clear test for added value: if sentiment contains predictive information about market tensions, demand expectations, or price prospects, it should improve forecasts of residual PP values compared to the baseline scenario based solely on the market.

The goal and the horizon. The goal is the resid_PP value for one month in advance. Forecasts are generated monthly with a one-month horizon. We evaluate performance under a rolling exit from the sample at the end of the sample (the same evaluation window that is used throughout the project), where each prediction is made using only the information available up to that month.

Rolling out-of-sample procedure. We use a backtest with a sliding start (an expanding window). At each stage, the model is re-evaluated based on all the data available up to month t, and a single forecast for month t+1 is created. This approach simulates the real workflow of forecasting and avoids overly optimistic conclusions based on sampling. For clarity of reporting, it is useful to think of this as a "window of learning that expands over time" with a test point one step ahead that shifts forward each month.

Model selection. All prediction specifications are intentionally simple linear regressions. This choice preserves the interpretability of the coefficients, reduces the risk of overfitting given the limited size of the monthly sample, and makes it easier to attribute any performance gain specifically to mood variables rather than model complexity.

	Model A (Baseline,	Model B (Sentiment	Model C (Sentiment

	market-only)	hybrid)	sections)
Target	resid_PP at t+1 (PP_EU residual vs crude baseline)	resid_PP at t+1	resid_PP at t+1
Autoregressive term(s)	resid_PP at t	resid_PP at t (and optionally t−1)	resid_PP at t (and optionally t−1)
Market regressors	Crude return at t−1; PGP return at t−1	Crude return at t−1; PGP return at t−1	Crude return at t−1; PGP return at t−1
Sentiment regressors	none	Overall OPEC sentiment index at t and t−1	Overall sentiment at t and t−1; plus section indices (Demand, Supply, Price Outlook) at t and t−1
Interpretation	predicts next-month PP residual using persistence + last month's market signals	tests whether the overall OPEC tone improves forecasts beyond market-only baseline	tests whether more granular sentiment components add incremental value beyond the hybrid index

Table 7 Forecasting models for next-month PP residual moves (1-month ahead)

Model specifications. We compare three nested models in which information about moods is gradually added. Model A is a basic model based solely on market data. Model B is complemented by the general (hybrid) OPEC sentiment index. Model C is complemented by more detailed sentiment data by including separate supply, demand, and price forecast indices. All models use lag structures corresponding to economic timing (market variables and past balances are introduced with lags; sentiment is being tested both simultaneously and with a one-month lag to account for the transmission delay in polymer prices).

Evaluation metrics. Efficiency is assessed using MAE and RMSE to measure the mean and quadratic error of the forecast, as well as directional accuracy, which measures how often the model correctly predicts the sign of the next month's residual value (i.e., whether PP exceeds or is inferior to the expected oil price movement). Directional accuracy is included because when making decisions about commodities, correctly predicting the direction of deviation can be important, even if the exact values are inaccurate.

8 Results and Model Comparison

This section presents the results of out-of-sample forecasting for the three nested models defined in Section 8 (A_baseline, B_sentiment_hybrid, C_sentiment_sections). The estimate is a rolling test 1 month ahead for the last 24 months (N = 24 forecasts), so each data point represents a genuine forecast for the "next month", made using only information from the past.

8.1 Forecast vs Actual behavior in the test period

Figure 8.1 reproduces the “Forecast vs Actual (test period)” plot comparing the three model forecasts against the realized target (PP residual).

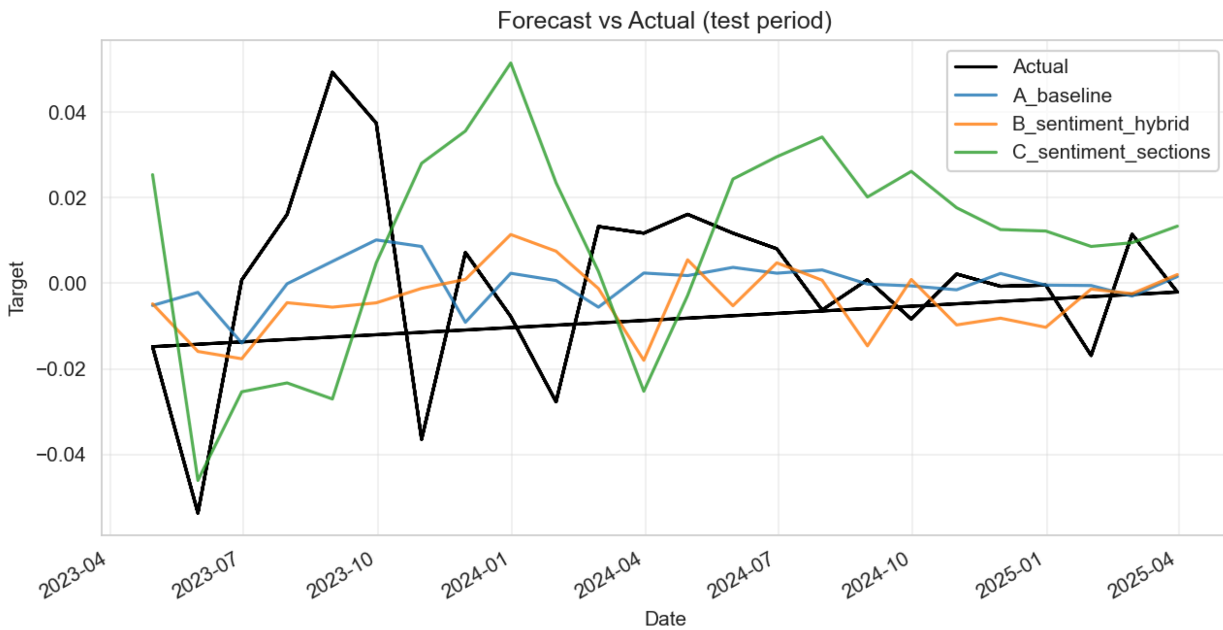

Figure 8.1 Forecast vs Actual plot

Visually, Model A tracks the general direction of the residual series better than the sentiment-augmented variants, while Models B and C tend to stay closer to zero and do not consistently match the larger swings in the realized series. Model C shows the largest deviations in several intervals, indicating that adding multiple section indices increases instability rather than improving tracking.

8.2 Forecast errors over time

Figure 8.2 should show the forecast errors (y_true − y_pred) for the three models over the same test months. The key pattern is that the error series for Models B and C does not become uniformly smaller than Model A; instead, the sentiment models frequently exhibit larger positive/negative error spikes. In particular, Model C has the widest error swings, consistent with overfitting/instability when adding many sentiment regressors on a short monthly sample.

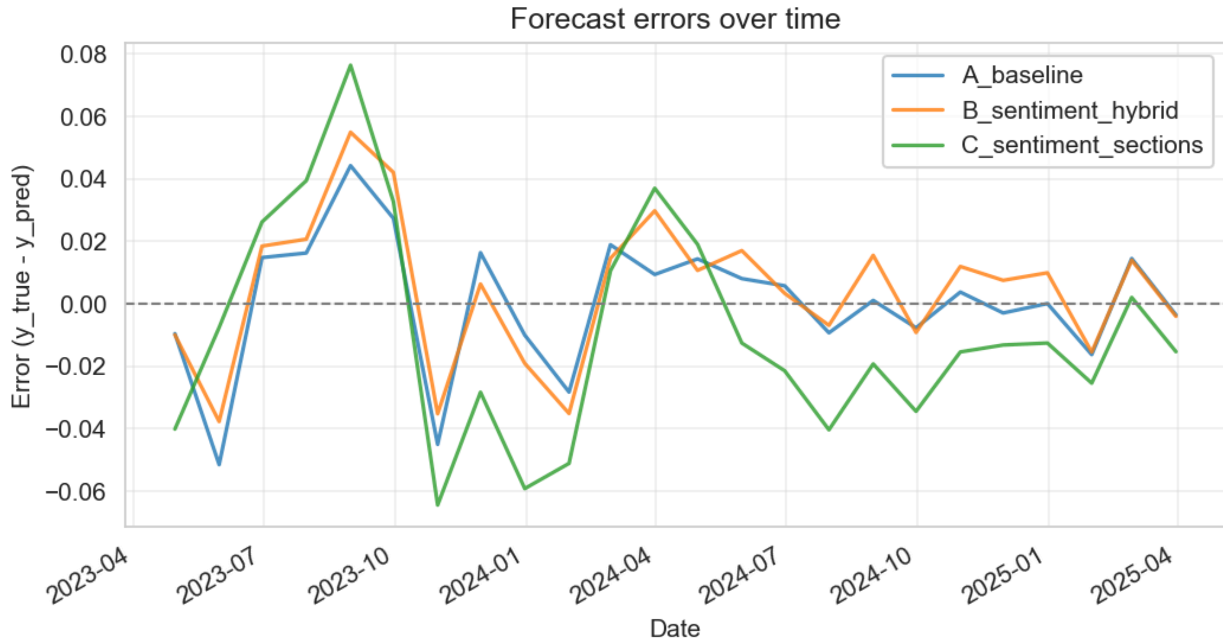

Figure 8.2 Forecast errors over time

8.3 Quantitative comparison of models (main metrics)

Table 8.3 reports the main evaluation metrics: MAE, RMSE, and Directional Accuracy. The best model should be highlighted (lowest MAE/RMSE and highest Directional Accuracy).

	Model A_baseline	Model B_sentiment_hybrid	Model C_sentiment_sections
MAE	0.0157	0.0187	0.0293
RMSE	0.0209	0.0228	0.0347
Directional Accuracy	0.4583	0.3750	0.3750

Table 8.3: Out-of-sample performance (N=24)

Across all three headline metrics, Model A performs best: it has the lowest MAE and RMSE, and it also achieves the highest directional accuracy. Adding the overall sentiment index (Model B) worsens error metrics and reduces directional accuracy. Adding section-level indices (Model C) further degrades performance substantially.

A note on MAPE: because the target is a residual series that can be near zero and can change sign, MAPE is not a reliable comparison metric here (it can behave erratically when actual values are close to zero). For that reason, MAE/RMSE and directional accuracy are treated as the main decision metrics.

8.4 Coefficient interpretation and relative importance

Figure 8.4 includes the "Model coefficients (full sample)" bar chart. This plot is useful for interpretability and for understanding why sentiment does not materially improve forecasts.

The baseline model (A) is primarily driven by the market variables and the autoregressive term: ret_PGP_lag1 appears as the largest-magnitude coefficient, while ret_CRUDE_lag1 is comparatively small in magnitude. The lagged residual term (resid_PP_lag1) is also included, capturing persistence in the residual component.

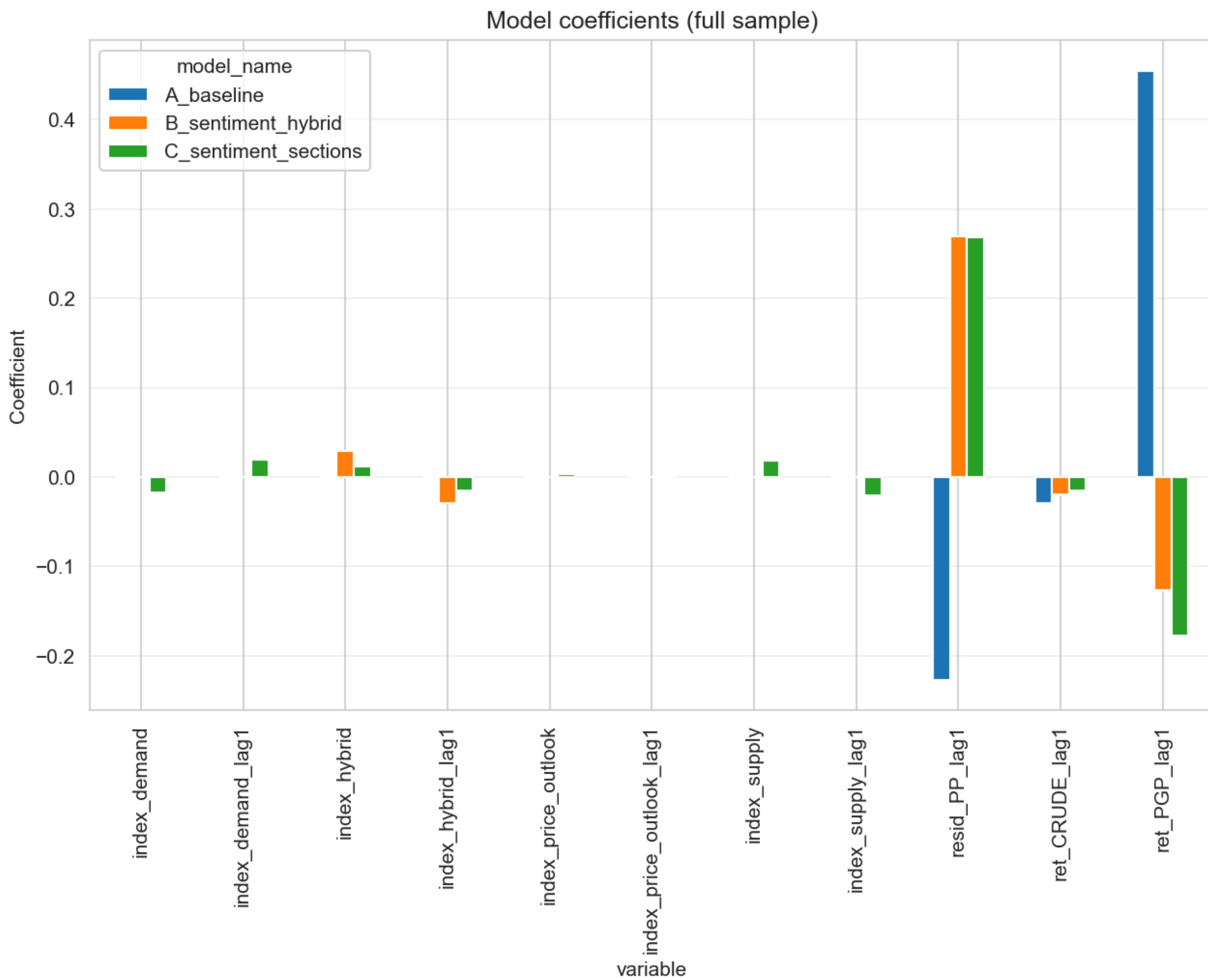

Figure 8.4 Model Coefficients (whole sample) bar chart

In the sentiment-augmented models (B and C), the sentiment coefficients (overall hybrid index and the demand/supply/price-outlook indices and their lags) are clustered close to zero relative to the market regressors. This indicates that, once the baseline information (past residual + crude and PGP returns) is included, the sentiment variables do not contribute a stable incremental signal in the linear specification.

The results provide a clear ranking: Model A (baseline) is the most effective model in the out-of-sample test, and increasing the number of signals does not improve prediction accuracy in this configuration. Model B slightly degrades accuracy and directivity, while Model C shows worse results, which corresponds to an increase in model complexity without a commensurate increase in the useful signal.

9 Conclusion

This paper looks at what drives polypropylene (PP) prices and how to forecast them better. Key findings show that PP prices connect to energy markets, but not fully. Correlations with crude benchmarks like Brent and WTI sit around 0.70 to 0.74. Natural gas links weaker at about 0.65. Normalized charts reveal shared big moves, like the 2020 crash and the 2021-2022 spike. PP and PGP track close, but PP stays smoother with smaller swings and stickier drops than crude.

A large part of PP changes stays unexplained by crude alone. Residuals after crude models show this idiosyncratic component. It comes from PP-specific factors, seasonal patterns, and unique shocks. Decomposition highlights a slow trend with modest yearly cycles and volatility that peaks in crises but stays lower for PP than crude.

The hybrid OPEC sentiment index ties weakly to PP residuals. Short lags of 1-3 months show small positive links, around 0.02 to 0.11. Changes in tone, like bullish shifts in 2020-21 or 2023-24, match some price turns. Yet OPEC reports often react to market events rather than lead them.

Adding sentiment to models helps a little on short-term forecasts. It cuts errors some against baselines like naive, Holt-Winters, or crude-OLS. But the gains stay small. Sentiment coefficients remain tiny, 0.00 to 0.04, next to stronger PGP or crude lags like 0.45. Even advanced tools like Random Forest add little over simple linear setups.

To answer the research questions directly:

- PP prices co-move moderately with crude and energy, but time-varying and imperfect.
- A big portion of movements—often short-term surprises—comes from idiosyncratic factors not captured by crude.
- OPEC sentiment links weakly to residuals, more as a reflector than a driver.
- Sentiment adds minor value to forecasts, but does not change performance much.

This project shows that combining NLP with commodity models has promise but limits here. Text from OPEC reports captures macro tones well. It describes ongoing conditions already in prices. Yet it adds little new signal for one-month PP predictions. Traditional market data like PGP and crude dominate. NLP works best as a small extra layer, not a main tool. For better results, try broader news sources, real-time text, or longer data spans. This highlights challenges in using sentiment for commodities tied tight to fundamentals. Still, the approach opens doors for hybrid models in volatile markets.

10 Appendix

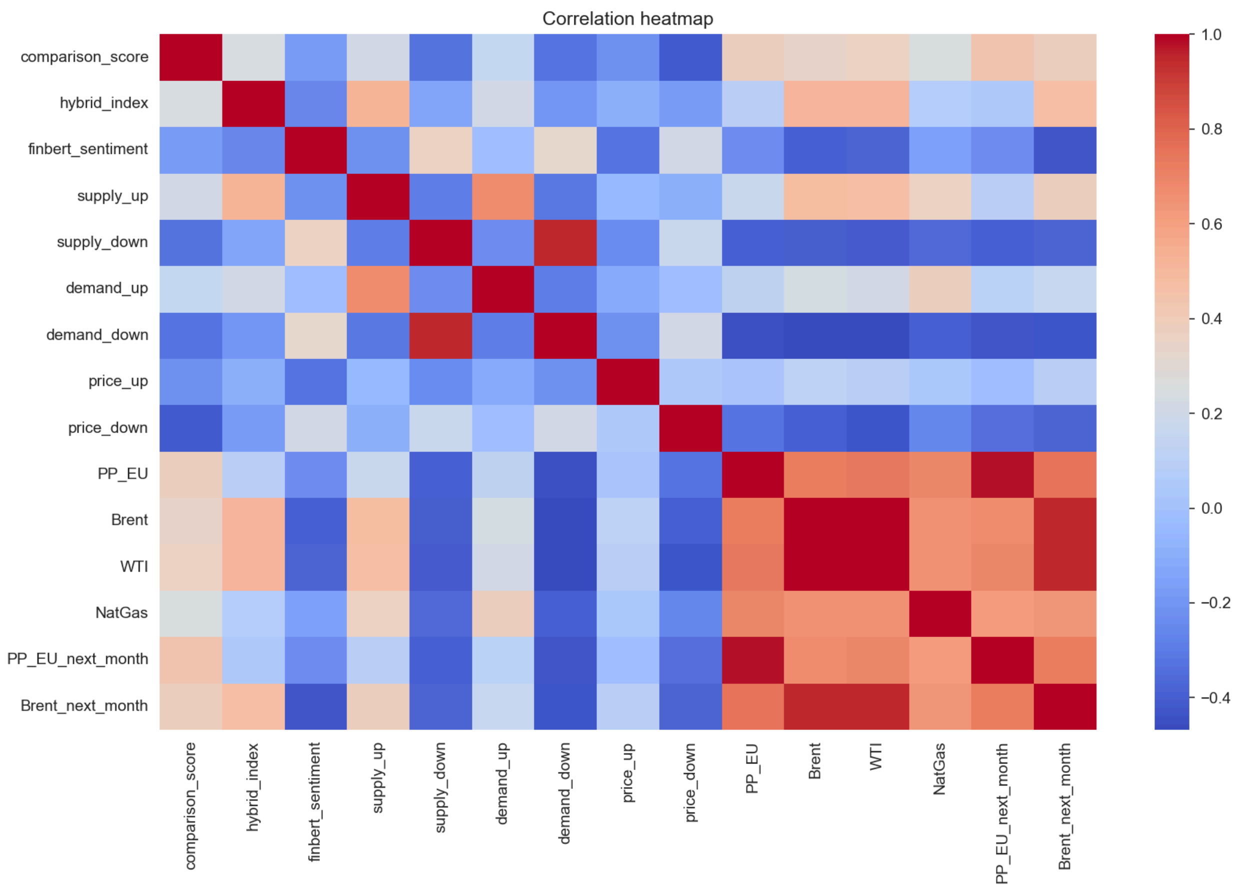

Figure 10.1 Correlation heatmap

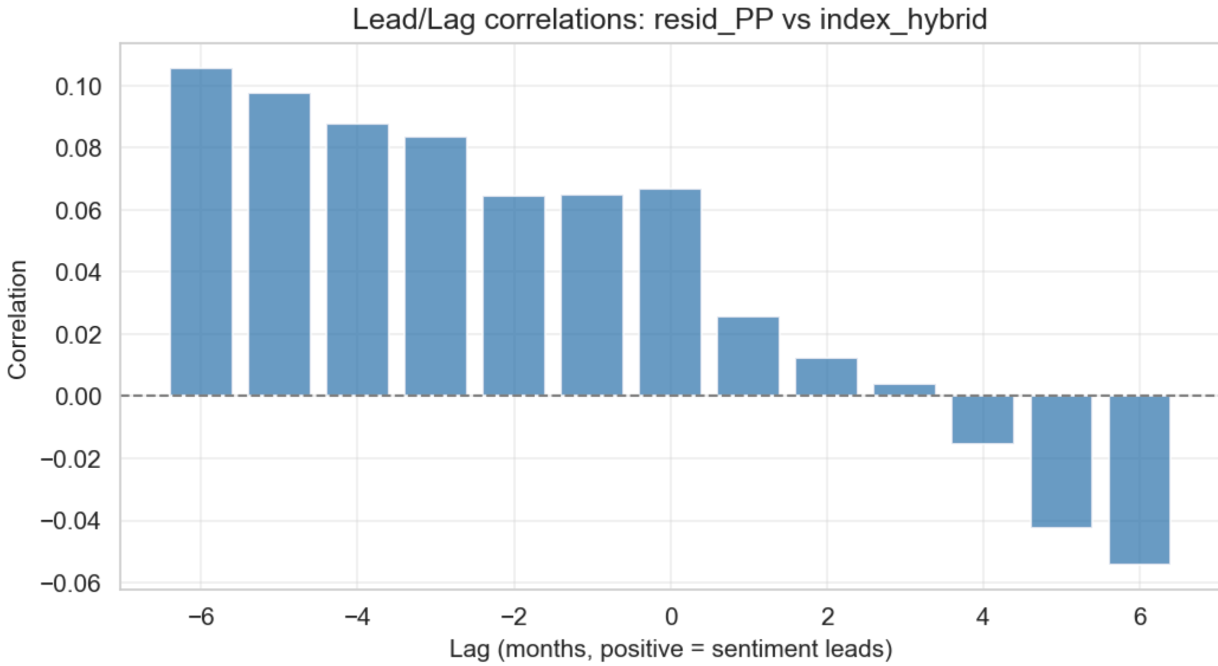

Figure 10.2 Lead/lag correlations bar plot

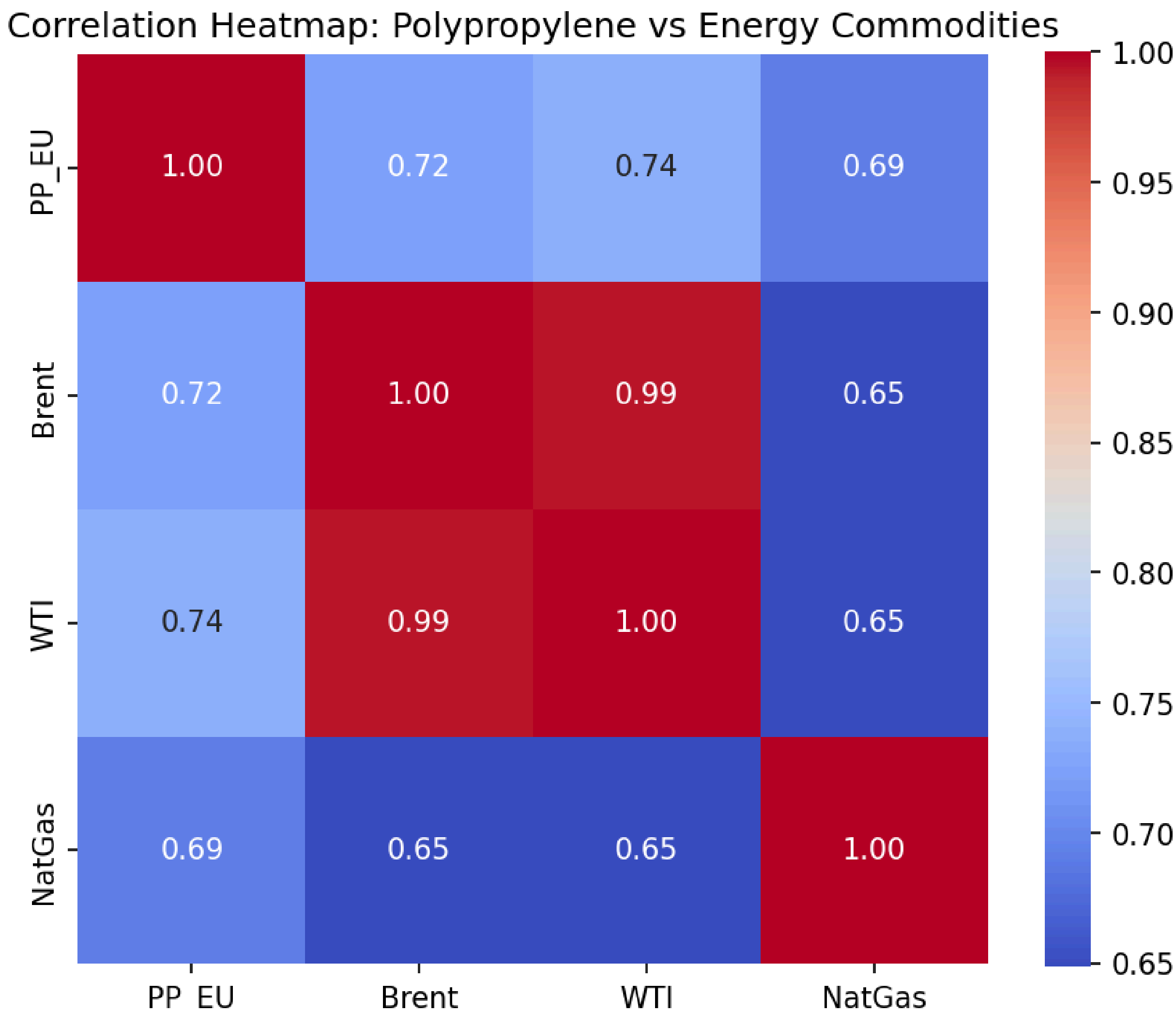

Figure 10.3 Correlation heatmap

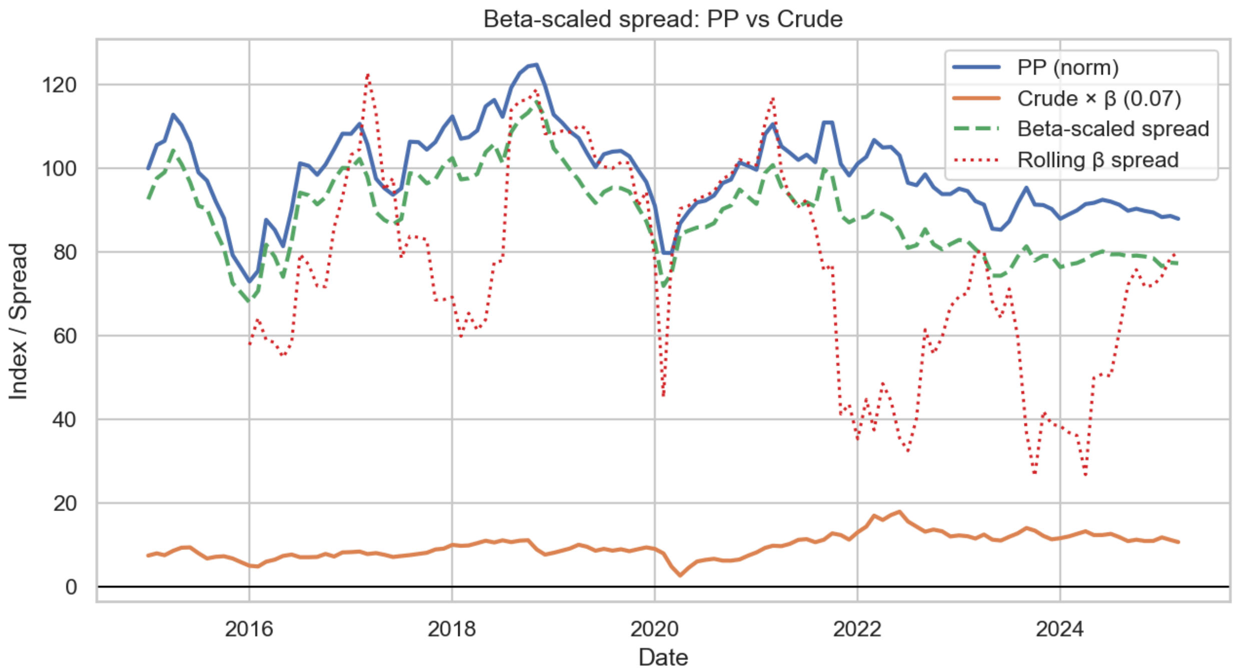

Figure 10.4 Beta-scaled spread: PP vs Crude oil

11 References

Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models (arXiv:1908.10063). arXiv. https://arxiv.org/abs/1908.10063

Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics & Data Analysis, 120, 70–83. https://doi.org/10.1016/j.csda.2017.11.003

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess (with discussion). Journal of Official Statistics, 6(1), 3–73. https://www.math.unm.edu/~lil/Stat581/STL.pdf

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Gardner, E. S., Jr. (1985). Exponential smoothing: The state of the art. Journal of Forecasting, 4(1), 1–28. https://doi.org/10.1002/for.3980040103

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts. https://otexts.com/fpp3/

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review, 99(3), 1053–1069. https://doi.org/10.1257/aer.99.3.1053

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Organization of the Petroleum Exporting Countries (OPEC). (n.d.). Monthly oil market report. Retrieved December 31, 2025, from https://www.opec.org/monthly-oil-market-report.html opec.org

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. International Journal of Forecasting, 16(4), 437–450. https://doi.org/10.1016/S0169-2070(00)00065-0